

Cuestionario de evaluación de las competencias docentes en el ámbito universitario. Evaluación de las competencias docentes en la universidad

Juan Antonio Moreno-Murcia^{1*}, Yolanda Silveira Torregrosa², Noelia Belando Pedreño³

¹Departamento de Psicología de la Salud, Universidad Miguel Hernández de Elche {j.moreno@umh.es}

²Departamento de Psicología de la Salud, Universidad Miguel Hernández de Elche
{silveiratorregrosa.y@gmail.com}

³Departamento de Psicología de la Salud, Universidad Miguel Hernández de Elche {nbelando@umh.es}

Recibido el 23 Octubre 2014; revisado el 24 Octubre 2014; aceptado el 10 Noviembre 2014; publicado el 15 Enero 2015

DOI: 10.7821/naer.2015.1.106

RESUMEN

El objetivo del estudio ha sido diseñar y validar un instrumento de medida para evaluar la intervención del docente universitario. Se administró el cuestionario de *Evaluación de la intervención Docente* (CEID) a 1297 universitarios. Se realizaron diferentes análisis factoriales (exploratorio y confirmatorio), de la consistencia interna, estadísticos descriptivos y de correlación de todos los ítems. Los datos obtenidos confirmaron una adecuada estructura psicométrica para la escala CEID, que quedó compuesta por tres dimensiones (planificación, desarrollo y resultado). Se concluye por tanto que es un instrumento válido y fiable para evaluar la intervención del docente universitario.

PALABRAS CLAVE: EVALUACIÓN DOCENTE, INSTRUMENTO DE MEDIDA, ESTIMACIÓN, DOCENCIA UNIVERSITARIA, COMPETENCIA DOCENTE.

1 INTRODUCCIÓN

En la actualidad, como consecuencia de la adaptación de los estudios al Espacio Europeo de Educación Superior (EEES), la formación universitaria requiere de una serie de modificaciones importantes respecto a la que se venía realizando hasta el momento. Esta adaptación exige a las universidades mejorar la calidad de sus actividades por medio de mecanismos de evaluación sistemáticos, tanto internos como externos, siendo uno de los pilares fundamentales de esta adaptación el apoyo a los procesos de innovación y desarrollo científico-técnico que supone una serie de cambios a nivel organizativo y pedagógico con los que se podría favorecer el interés hacia métodos pedagógicos innovadores (Marqués, 2008). En este proceso de renovación metodológica de las universidades, la valoración de los estudiantes podría ser determinante para conocer e intervenir sobre la calidad formativa de los docentes, por lo que sería necesario diseñar instrumentos de medida en los que se valore la dinámica del proceso de enseñanza-aprendizaje atendiendo a la metodología y competencias docentes.

Uno de los cambios más importantes en el nuevo modelo es la formulación de las nuevas titulaciones a partir de las competencias profesionales (Carreras, 2005). Este enfoque por competencias supera el tradicional sistema de formación universitaria centrado exclusivamente en materias o disciplinas. La formación basada en competencias profesionales pretende dotar a los aprendizajes de significatividad y funcionalidad y poner en acción una serie de conocimientos, actitudes, habilidades y valores en el desarrollo de la actividad profesional, implicando al estudiante de manera activa, favoreciendo así la reflexión y el pensamiento crítico, permitiéndole tomar decisiones, resolver problemas, diseñar proyectos, etc. Es decir, la utilización de estilos autorregulados que promueven la autonomía facilitarían la integración de los conocimientos que va adquiriendo (Leary, Walker, Shelton, & Harrison, 2013) llevando al estudiante a mostrar un mayor interés y compromiso por la materia que está estudiando y mejorar su rendimiento académico (Vansteenkiste, Soenens, Sierens, Luyckx, & Lens, 2009). Este nuevo modelo supone un cambio importante en la manera de concebir el proceso de enseñanza-aprendizaje en las universidades, en el que el docente no solo debe diseñar y presentar los contenidos a desarrollar. Se trata de incentivar la implicación docente y desarrollo del conocimiento crítico de los estudiantes, para lo que se deberían concretar las metodologías de enseñanza y los sistemas de evaluación a desarrollar para cada materia, estando en relación con las competencias profesionales previamente establecidas facilitando así su transferencia al ámbito profesional (Salmerón, 2013). Según este modelo de formación basado en competencias, los sistemas de evaluación y los métodos docentes se estructuran y diseñan en función a las competencias a alcanzar, de modo que trabajar en el nuevo modelo universitario va a suponer una mayor exigencia por parte del profesorado (Palliser, Fullana, Planas, & del Valle, 2010). Incorporar las dinámicas de evaluación interna en la gestión de las titulaciones, va a requerir especialmente de un elevado grado de motivación en el profesorado. En este sentido, la evaluación de la actividad docente parece adquirir un valor significativo dentro de los nuevos sistemas de garantía de la calidad de las instituciones universitarias, de manera que podemos entender estos procesos de evaluación como un medio

*Por correo postal, dirigirse a:
Universidad Miguel Hernández de Elche, Centro de Investigación del Deporte
Avenida de la Universidad, s/n. - 03202 Elche (Alicante)

de reflexión para poner en marcha políticas de acción y mejorar la calidad de la intervención docente. En este cometido, las instituciones universitarias, para alcanzar una mejora de la calidad en el modo en el que es afirmado por la European Association for Quality Assurance in Higher Education (ENQA), deberían desarrollar y disponer de mecanismos formales como la aprobación, la revisión periódica y el control de sus programas y títulos, a través de sistemas de control como la evaluación objetiva.

En España encontramos algunos antecedentes de la evaluación de la calidad en la educación universitaria como el Programa de Evaluación de la Calidad Docente (1980), el Programa Experimental de Evaluación de la Calidad del Sistema Universitario (1993,1994) y a partir 1996 las universidades han evaluando la calidad de sus titulaciones a través de programas como el Plan Nacional de Evaluación de la Calidad de las Universidades, el II Plan de Calidad de las Universidades, el Programa de Evaluación Institucional de ANECAⁱ (2007) o mediante otros programas similares autonómicos (ACSUGⁱⁱ, AVAPⁱⁱⁱ ó AQU^{iv}). En los últimos 15 años las universidades han iniciado una evaluación de sus servicios apoyados en enfoques centrados en la gestión de la calidad. El 21 de diciembre de 2006 el Ministerio de Educación y Ciencia (MEC) publicó las *Directrices para la elaboración de títulos universitarios de grado y máster*, donde se incluyó la necesidad de introducir un Sistema de Garantía de la Calidad como uno de los elementos imprescindibles a la hora de proponer títulos en el futuro. Hasta no hace mucho las instituciones universitarias utilizaban un sistema de evaluación basado prácticamente en la actividad investigadora, dejando en un segundo plano la labor educativa del docente. Es decir, se valoraba al docente atendiendo fundamentalmente al número total de publicaciones que realizaba, al índice de impacto de donde se recogía la publicación, y a las veces que se citaba un determinado artículo. Asimismo, la evaluación docente se centraba únicamente en valorar la adquisición o no de conocimientos por parte de los estudiantes en términos de mayor o menor rendimiento académico (Buendía, 1996). Y son los estudiantes los que con mayor frecuencia han sido escogidos para evaluar la calidad de la intervención docente. Al respecto, Tejedor (1996) apunta que el criterio de referencia de los estudiantes no debe de ser el único, ya que existen numerosas aportaciones en relación a los agentes de evaluación del docente, siendo los principales agentes los estudiantes, otros docentes y el propio docente (García y Congosto, 2000). De esta manera, la evaluación docente ha ido evolucionando de tal forma que en la actualidad ha sido considerada como un elemento importantísimo para analizar la calidad formativa y profesional de las instituciones educativas. Esta nueva perspectiva establece modelos de evaluación que parten de la autoevaluación como un proceso reflexivo y participativo, con la intención de implicar a todos los miembros del sistema de cara a ofrecer una mayor calidad.

En relación con los instrumentos utilizados para evaluar la intervención docente, se ha observado una cierta disparidad entre las universidades españolas, aunque presentan ciertos aspectos en común. Se han realizado una serie de estudios que han tratado de analizar los instrumentos empleados en cada una de las universidades españolas. Por ejemplo, Muñoz, Ríos de Deus, y Abalde (2002) realizaron una revisión sobre los instrumentos y dimensiones utilizados en las distintas universidades y diseñaron un instrumento nuevo de 40 ítems

agrupados en 10 dimensiones (cumplimiento con las obligaciones, infraestructura, programa, conocimiento-interrelación con la materia, metodología, materiales, actitud del docente, evaluación, prácticas y satisfacción). En esta misma línea, López-Barajas, y Carrascosa (2005) elaboraron un cuestionario compuesto por 25 ítems desglosados en cuatro dimensiones: interacción con el alumnado, metodología, obligaciones docentes y evaluación, medios y recursos, para evaluar la docencia en la Universidad de Jaén. Valiéndose del criterio del alumnado determinaron cuáles eran las dimensiones más valoradas y las que más predecían la valoración de la docencia, siendo la dimensión “interacción con el alumnado la que predijo en mayor medida la valoración global del docente. Por otro lado, Casero (2008) elaboró un instrumento formado por 92 ítems agrupados en infraestructura, programa, conocimiento/interrelación de la materia, metodología, materiales, actitud del profesor, evaluación, prácticas y satisfacción para evaluar a los docentes universitarios, buscando consenso entre estudiantes y docentes respecto a la importancia que le otorgaban a cada uno de los ítems con el objetivo de crear el instrumento ideal para la evaluación docente. Asimismo, González y López (2010) trataron de analizar cuáles eran las competencias docentes que las universidades españolas consideran esenciales para definir el rol del profesorado y establecer así un sistema de indicadores de evaluación de la calidad docente, elaborando un instrumento piloto de evaluación formado por 40 ítems agrupados en las siguientes dimensiones: diseño de los programas/guías docentes de la asignatura, metodología docente, actividades prácticas, recursos didácticos, planificación docente, sistemas de evaluación, cumplimiento de las obligaciones docentes y actitud del profesorado, donde a diferencia del estudio de Casero (2008), la creación del perfil competencial del profesorado se obtuvo a través de un proceso de valoración de expertos y no por consenso entre el profesorado y alumnado.

Diversos docentes de la Universidad Politécnica y Autónoma de Madrid realizaron un estudio en el 2010, donde pretendían desarrollar un sistema de indicadores de calidad para la evaluación de la actividad docente universitaria a partir del análisis de los cuestionarios empleados en las diferentes universidades españolas. Aunque las dimensiones propuestas por las diferentes universidades oscilan entre 2 y 9, en la mayoría de ellas se agrupan los ítems según las tres dimensiones del modelo propuesto por ANECA (planificación, desarrollo y resultado), a las cuales se les añade alguna otra dimensión como es el entorno docente o la satisfacción. La Universidad de Navarra emplea un cuestionario similar formado por 24 ítems agrupados en cuatro dimensiones (planificación, desarrollo, resultados y satisfacción global). Análogamente, González y López (2010) realizaron un estudio en el que trataron de estimar el nivel de validez del Modelo MIECEES^v, así como el grado de importancia que otorgaban los estudiantes a cada uno de los ítems para conseguir una universidad de calidad. El instrumento finalmente quedó compuesto por seis dimensiones, 19 indicadores y 85 ítems que definían operativamente los indicadores del modelo de evaluación de la calidad en el Espacio Europeo de Educación Superior.

Pozo, Bretones, Martos, y Alonso, (2011) realizaron una revisión por comunidades autónomas de las universidades españolas que utilizan el modelo DOCENTIA para evaluar la actividad de sus docentes universitarios, donde explican la

adaptación de este modelo “Docentia-Almería” realizada por Pozo, Giménez, y Bretones (2009), con la finalidad de crear un marco común en relación a la evaluación del profesorado en las universidades andaluzas, estableciendo como principal diferencia a la propuesta de ANECA, una cuarta dimensión denominada “innovación y mejora”, en la que incluyen elementos asociados a los procesos de cambio incorporados por el docente que puedan contribuir a una mejor calidad de la enseñanza.

Aunque muchas universidades se apoyan en un mismo modelo, parece que existe gran diversidad a la hora de establecer qué factores son representativos de una metodología docente de calidad en las universidades. Así pues, considerando todos estos aspectos, el objetivo del estudio ha sido diseñar y validar un instrumento de medida que permita evaluar la intervención del docente universitario y para ello se espera dar validez y fiabilidad a un instrumento compuesto por tres dimensiones: planificación, desarrollo y resultado.

2 MÉTODO

2.1 Participantes

La muestra estuvo compuesta por 1297 estudiantes universitarios (762 hombres y 535 mujeres) con edades comprendidas entre los 18 y 27 años ($M = 24.75$; $DT = 5.09$), estudiantes de 1º a 4º grado. La muestra fue desglosada en dos submuestras, la submuestra 1 integrada por 649 estudiantes (410 hombres y 239 mujeres) utilizada para llevar a cabo el análisis factorial exploratorio y la submuestra 2 formada por 648 estudiantes (411 hombres y 237 mujeres) empleada en el análisis factorial confirmatorio.

2.2 Medidas

Evaluación docente. La versión final del *Cuestionario de Evaluación de Intervención Docente* (CEID) estaba formado por 28 ítems agrupados en tres factores: planificación con cuatro ítems (e.g. “Informa con claridad en el programa de los objetivos, bibliografía tutorías, contenidos y método de evaluación”), desarrollo con 17 ítems (e.g. “Propicia el trabajo en equipo”) y resultado con siete ítems (e.g. “Aplica los criterios de evaluación de las actividades correspondientes a los establecidos en el programa”). Los ítems son valorados a través de una escala tipo Likert con un rango de puntuación de 1 (*Totalmente en desacuerdo*) a 5 (*Totalmente de acuerdo*) encabezados por la frase “El docente...”.

2.3 Procedimiento

La construcción del cuestionario se realizó a partir de una extensa revisión bibliográfica donde se fue construyendo los aspectos más importantes a evaluar y que serían reflejados en los distintos ítems. La redacción de los mismos surgió del trabajo de campo de diferentes docentes universitarios expertos en la temática. Se elaboró una primera versión compuesta por 68 ítems, que tras ser revisada por expertos, quedó reducida a 40 ítems agrupados en tres factores: planificación, desarrollo y resultado. Además, la escala se sometió a una prueba de comprensión general donde se pidió a los estudiantes que leyeran detenidamente los ítems y que valoraran la importancia que tenía para ellos.

Se contactó con los responsables de grado, así como, con los coordinadores de curso para informarles del objetivo de la investigación. Los cuestionarios fueron cumplimentados en las

aulas antes del inicio de una de las clases teóricas en presencia del investigador principal y en ausencia del docente al que evaluaban. A cada asignatura y a su docente se le asignó un número de identificación. El investigador principal informó a los estudiantes del objetivo del estudio y de la correcta cumplimentación del cuestionario; además, resolvió las dudas que pudieran surgir, siendo necesario un tiempo aproximado de 10 minutos. La participación fue voluntaria y se preservó el anonimato de los participantes.

2.4 Análisis de datos

Para determinar y evaluar la estructura factorial del instrumento se llevaron a cabo un análisis factorial exploratorio y un análisis factorial confirmatorio, respectivamente. Asimismo, se analizó la consistencia interna del instrumento mediante el coeficiente de alfa de Cronbach y se obtuvieron los estadísticos descriptivos (medias y desviaciones típicas) y correlaciones bivariadas de todos los ítems. Para el análisis de los datos se utilizaron los paquetes estadísticos SPSS 21.0 y AMOS 21.0

3 RESULTADOS

3.1 Análisis factorial exploratorio

Tras varios análisis factoriales exploratorios se fueron eliminando algunos ítems por no llegar a una rotación mínima de .40. Finalmente, se llevó a cabo un análisis factorial de componentes principales con rotación oblimin directo con la submuestra 1 de los que resultaron 28 ítems que se agruparon en tres factores: planificación (formado por cuatro ítems: 4, 20, 21 y 22), desarrollo (formado por 17 ítems: 1, 2, 3, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 23 y 26) y resultado (formado por siete ítems: 5, 6, 14, 24, 25, 27 y 28). Se obtuvieron autovalores mayores de 1.00 (1.67, 1.50, y 1.34, respectivamente), explicando una varianza total del 68.2% (38.09%, 15.36% y 14.77%, respectivamente) (Tabla 1).

Tabla 1. Cuestionario de evaluación de la intervención Docente (CEID)

Ítems	Factor 1	Factor 2	Factor 3
1. Presenta los contenidos mínimos de su materia ajustándose al nivel de conocimientos del alumnado	-	.728	-
2. Es fácilmente accesible (tutorías, correos, etc.)	-	.733	-
3. Permite al estudiante la organización y distribución de parte de las tareas a realizar en la asignatura	-	.736	-
4. Informa con claridad en el programa de los objetivos, bibliografía, tutorías, contenidos y métodos de evaluación	.603	-	-
5. Informa de las competencias previstas que el estudiante tiene que adquirir	-	-	.569
6. Me proporciona información científica que me permite conocer más y mejor la materia	-	-	.694
7. Presenta los contenidos siguiendo un esquema claro y lógico, destacando lo importante	-	.605	-
8. Permite y estimula la participación del alumnado	-	-.757	-
9. Propicia el trabajo individual	-	-.828	-
10. Propicia el trabajo en equipo	-	-.834	-
11. Relaciona los aprendizajes con el ámbito profesional	-	-.577	-
12. Realiza síntesis iniciales y finales de la sesión y/o tema en clase	-	-.632	-

Tabla 1 (continuación)

13. Favorece el interés del alumnado y la motivación por aprender	-	-.614	-
14. Fomenta la investigación y el espíritu crítico de los estudiantes	-	-	.530
15. Facilita la interacción estudiante-estudiante y estudiante-docente	-	-.712	-
16. Atiende y responde con claridad las consultas realizadas en clase	-	.719	-
17. Atiende adecuadamente las tutorías que se le solicitan	-	.737	-
18. Mantiene una postura objetiva y respetuosa con el alumnado	-	.577	-
19. Organiza actividades para que el estudiante participe activamente en las tareas de clase	-	-.653	-
20. Diseña y relaciona los contenidos teóricos con los prácticos	.649	-	-
21. Incorpora y emplea con eficiencia las TICs (Tecnologías de la Información y la Comunicación)	.498	-	-
22. Domina los contenidos de la asignatura	.692	-	-
23. Interrelaciona los contenidos de la materia con otras asignaturas	-	.598	-
24. Aplica el programa establecido con cierta flexibilidad, para una mejor dinámica de clase	-	-	.625
25. Utiliza recursos materiales que facilitan el aprendizaje	-	-	.602
26. Se relaciona satisfactoriamente con el alumnado	-	-.609	-
27. Diseña el contenido y desarrolla la asignatura para favorecer la adquisición de competencias profesionales	-	-	.591
28. Aplica los criterios de evaluación de las actividades correspondientes a los establecidos en el programa	-	-	.468
Varianza explicada de cada factor	38.09%	15.36%	14.77%
Varianza total explicada	68.20%		

3.2 Análisis factorial confirmatorio

Se realizó un análisis factorial confirmatorio (CFA) con la submuestra 2, basándonos en 28 medidas observadas y en los tres constructos latentes que correlacionaban libremente (Anderson & Gerbing, 1988). El modelo fue reducido para mantener unos grados de libertad razonables (Cecchini, González, & Montero, 2007; Ntoumanis, 2001). Así, los ítems que componían cada una de las dimensiones fueron divididos en tres grupos (Marsh, Richards, Johnson, Roche, & Tremayne, 1994). La validez del modelo de medición fue considerada a través de una serie de coeficientes fit, también llamados índices de bondad de ajuste: χ^2 , $\chi^2/\text{d.f.}$, RMSEA y los índices incrementales (CFI, NFI y TLI). Se utilizó el método de estimación de máxima verosimilitud junto con el procedimiento de *bootstrapping*, ya que el resultado del coeficiente multivariado de Mardia fue 61.56, lo que indicaba falta de normalidad multivariada de los datos. Por ello, siguiendo a Finney y DiStefano (2006) se utilizó el método robusto de estimación de máxima verosimilitud (Byrne, 2001). Tras un primer análisis, los resultados globales del modelo indicaron un ajuste moderado ($\chi^2 (59, N = 649) = 1899.01, p = .000; \chi^2/\text{d.f.} = 5.47; \text{CFI} = .80; \text{NFI} = .77; \text{TLI} = .81; \text{RSMR} = .05; \text{RMSEA} = .08$). Con el objeto de mejorar el ajuste, se procedió a realizar un análisis del test de Langrange (LMtest), mostrando que se producía un incremento en el ajuste si se correlacionaban los errores (9 interacciones) de determinados ítems que pertenecían a un mismo factor y cuyo nivel de significación era $p <$

.000., todos pertenecientes al factor desarrollo. Los índices obtenidos fueron adecuados: $\chi^2 (69, N = 649) = 1051.03, p = .000; \chi^2/\text{d.f.} = 3.12; \text{CFI} = .91; \text{NFI} = .90; \text{TLI} = .90; \text{RSMR} = .06; \text{RMSEA} = .06$.

3.3 Análisis de consistencia interna

La consistencia interna de cada uno de los factores resultantes del análisis factorial (alfa de Cronbach) presentó los siguientes resultados: .70 para la planificación, .91 para el desarrollo y .79 para el resultado. Considerando todos los ítems en un solo factor la consistencia obtenida fue de .94.

3.4 Análisis descriptivo y correlaciones bivariadas

Los valores medios para los factores planificación, desarrollo y resultado fueron, 3.95, 3.77, y 3.74, respectivamente. Los datos del análisis de correlación revelaron que los tres factores correlacionaron positivamente entre sí (Tabla 2).

Tabla 2. Estadísticos Descriptivos y Correlaciones de los cuatro factores del cuestionario

Variables	M	DT	1	2	3
1. Planificación	3.95	.64	-	.72**	.73**
2. Desarrollo	3.77	.57	-	-	.82**
3. Resultado	3.74	.58	-	-	-

Nota: ** $p < .001$; M = Media; DT = Desviación típica.

4 DISCUSIÓN

Según Arranz (2007), cada vez más la sociedad reclama un funcionamiento de las universidades vinculado a criterios de eficacia, eficiencia y excelencia. Con lo que el propósito de este estudio fue diseñar y validar un instrumento de medida que permita evaluar la intervención del docente universitario. En este sentido, parece importante contar con procesos de evaluación de las metodologías de actuación que respondan a las necesidades educativas y profesionales implantadas por el EEES. Así, para que estos procesos de evaluación sean exitosos, podría ser fundamental una formación en el docente universitario en relación a técnicas y procedimientos para que participe de manera más activa en los sistemas de evaluación. Para ello, el diseño y validación de instrumentos de medida podrían ser una herramienta clave en el intento de optimizar el funcionamiento de la intervención docente y poder así realizar una valoración objetiva de los resultados (García y Congosto, 2000). Con este propósito se diseñó el cuestionario de *Evaluación de la Intervención Docente*, donde las sugerencias aportadas por los expertos y los respectivos análisis factoriales confirmatorios mostraron un instrumento compuesto por 28 ítems agrupados en tres factores con una adecuada consistencia interna (planificación, desarrollo, y resultado). Donde la planificación de la docencia hace referencia a todo lo relacionado con el proceso previo de reflexión y diseño de la materia (la programación, la organización de clases teóricas, prácticas, tutorías, planificación de actividades de aprendizaje previstas, criterios y métodos de evaluación, materiales y recursos para la docencia). El desarrollo de la enseñanza está relacionado con todo lo relativo a la puesta en práctica y cumplimiento del programa, la realización de actividades de enseñanza y aprendizaje realizadas, actividades pedagógicas previstas, así como procedimientos de evaluación aplicados. Y los resultados,

en términos de objetivos formativos, hacen referencia a los logros alcanzados por los estudiantes, y a los aspectos relacionados con la revisión y mejora de la actividad docente, reconocimiento externo de la labor docente y creación de materiales docentes. Aunque encontramos gran diversidad a la hora de determinar cuáles son los principales factores a evaluar en la práctica docente, la mayoría de universidades españolas toma como referencia el modelo propuesto por ANECA (Pozo, Bretones, Martos, & Alonso, 2011). En esta línea, Pozo et al. (2009) adaptan este modelo para la Comunidad de Andalucía, agrupando las dimensiones en planificación, desarrollo, resultados e innovación y mejora. Otras universidades optan por proponer un mayor número de dimensiones siendo éstas más específicas, como el diseño de la programación, la metodología docente, las actividades prácticas, los recursos didácticos, los sistemas de evaluación y la actitud del profesorado (González-López & López-Cámara, 2010). Igualmente agrupan las dimensiones en criterios de interacción con el alumnado, metodología, obligaciones docentes y evaluación, medios y recursos (López-Barajas y Carrascosa, 2005). Por otro lado, Casero (2008) para evaluar la intervención del docente, engloba los factores en infraestructura, programa, conocimiento/interrelación de la materia, metodología, materiales, actitud del profesor, evaluación, prácticas y satisfacción del alumnado. Mientras que, González y López (2010) las agrupan en competencias, información universitaria, rol del estudiante universitario, sistemas de orientación y tutoría, metodología y satisfacción del alumnado. Lukas, Karlos, Etzeberria, y Lizasoain (2014) proponen un instrumento para evaluar la actuación docente formado por 4 dimensiones: Planificación de la docencia (5 ítems), Desarrollo de la docencia e intervención docente (11 ítems), Resultado (2 ítems) y además añaden 2 ítems de carácter abierto (aspectos positivos del docente y aspectos a mejorar). Mientras que, autores como Bol-Arreba, Sáiz-Manzanares, y Pérez-Mateos (2013) optan por adaptar el cuestionario Student Evaluation of Educational Quality (SEEQ) short version (Marsh, 1987). La nueva versión creada está formada por diez ítems (el interés del alumno por la asignatura, la utilización de los materiales docentes, el interés del profesorado hacia la asignatura, la accesibilidad del profesor en las horas de tutoría y fuera de clase, la utilidad de los materiales docentes, las pruebas de evaluación continua y su utilidad, el ajuste de los procedimientos de evaluación, la calidad y explicación de los materiales docentes, la motivación por parte del profesor a la participación del alumno en clase y la carga de trabajo de la asignatura).

Tenemos que la evaluación de la enseñanza realizado por los estudiantes (Students' evaluations of teaching" SET) es utilizada en la mayoría de universidades del todo el mundo como una medida de calidad docente, de forma que los resultados de las valoraciones ofrecen un feedback al docente para mejorar su enseñanza (Balam & Shannon, 2010). Según el trabajo de Kuzmanovic, Savic, Gusavac, Makajic-Nikolic, & Panic, (2013) los indicadores de una buena docencia y de calidad según los estudiantes son: la disponibilidad del docente, su claridad en las presentaciones, el enfoque metodológico y sistemático, la información proporcionada a los estudiantes sobre el trabajo de la asignatura, favorecer la participación activa en la clase, así como considerar y responder a las preguntas de los estudiantes. De la misma manera, Shukrie (2011) determina que los docentes más eficaces son aquellos que conocen las necesidades de sus estudiantes y tratan de aumentar su motivación, emplean diferentes estrategias de enseñanza, tienen buenas habilidades

comunicativas y promueven a los estudiantes a ser optimistas sobre sus capacidades.

Las universidades del todo el mundo han ido desarrollando complejos procedimientos e instrumentos para evaluar y garantizar la calidad educativa en la enseñanza superior a partir de las opiniones de sus estudiantes. Por ello nos encontramos numerosos instrumentos como son el "Students' Evaluation of Teaching Effectiveness Rating Scale" (Toland & De Ayala, 2005), el "Teaching Proficiency Item Pool" (Barnes et al., 2008), the questionnaire for student evaluation of teaching (SET 37, Mortelmans, & Spooren, 2009), the Exemplary Teacher Course Questionnaire (Kember & Leung, 2008), o el "Teaching Behavior Checklist" (Keeley, Furr, & Buskist, 2010). Y aunque existe cierta variedad de dimensiones o ítems de unos a otros para evaluar la docencia, en general las dimensiones identificadas son similares. En especial, las referentes a la organización, evaluación, y la relación docente-estudiante parecen ser las dimensiones más comunes en la mayoría de los instrumentos. Por ejemplo, Dodeen (2013) emplea un cuestionario de 29 ítems agrupados en cinco dimensiones: Conocimiento y organización del docente (7 ítems), explicación clara (6 ítems), calificación y evaluación (6 ítems), métodos de enseñanza (4 ítems), y la relación con los estudiantes (6 ítems). Toland y Alyala (2005) identificada tres dimensiones: entrega de la información del curso, interacción docente-estudiante, y la regulación del aprendizaje de los estudiantes. Para Gursoy y Umbreit (2005) son cuatro las dimensiones: aprendizaje, organización, la instrucción y la carga de trabajo.

Aunque estos cuestionarios presentan la validez y fiabilidad necesaria para evaluar la efectividad de la enseñanza (Khong, 2014; Onwuegbuzie, Daniels, & Collins, 2009), el trabajo de revisión realizado por Spooren, Brockx y Mortelmans (2013) advierte de que existen todavía ciertas controversias al respecto en la literatura, dado que la mayoría de los cuestionarios de evaluación de la enseñanza no han presentado suficientes pruebas de validez.

En el contexto español, a pesar de la diversidad de taxonomías sobre las dimensiones a evaluar, todas ellas presentan varios elementos en común y además han mostrado ser instrumentos válidos y fiables para la evaluación de la actuación docente, aunque el hecho de que un instrumento sea adecuado, es condición necesaria pero no es suficiente para garantizar una evaluación exitosa (Muñiz & Bartram, 2007).

Con el diseño de esta escala se podría contar con un instrumento que permite valorar la intervención docente del profesorado universitario, de forma que los resultados obtenidos en la evaluación podrán ser de gran utilidad para el docente, proporcionando datos objetivos. Por tanto, es fundamental que se consideren tanto los aspectos relacionados con el proceso de aplicación del instrumento, como el uso que se hace de los resultados del mismo, siendo importante que tras la evaluación haya un feed-back hacia las partes implicadas (Hattie, 2009; Leung, Jiang, & Busserb, 2013; Muñiz & Fonseca-Pedrero, 2008) y se entienda ésta como una estrategia para mejorar la enseñanza, que implique en la misma a los estudiantes, permitiendo así orientar la actividad docente (Padilla & Gil, 2008) y como consecuencia se lleve a cabo una serie de planes de mejora y cambios en su práctica pedagógica, que redefinan el proceso de enseñanza-aprendizaje que hasta ahora se venía desarrollando (López Ruiz, 2011) y que optimicen la calidad educativa universitaria, consiguiendo así una evaluación más formativa (Moore, Walsh, & Rísquez, 2012). Pero todo ello requiere de un cambio de actitud, viéndose reflejado en una

transformación de los estilos docentes (Gargallo, 2008). Todos estos cambios pedagógicos suponen una gran exigencia tanto para los docentes (Pallisera et al., 2010) como para los estudiantes implicados en el proceso.

5 CONCLUSIONES

Tras los diferentes análisis realizados, se ha estimado un instrumento de medida que presenta unas adecuadas propiedades psicométricas que le dan validez para la valoración de la intervención del docente universitario. Las tres dimensiones evaluadas (planificación, desarrollo y resultado) presentan una buena consistencia interna (.70, .91, .79 respectivamente), lo que garantiza la fiabilidad del instrumento. No obstante, este estudio presenta algunas limitaciones en cuanto al carácter integral del proceso de evaluación, que en este caso se centra exclusivamente en la valoración de los estudiantes y, aunque esta valoración es necesaria, no suficiente (Berk, 2013). Por ello serían necesarios otros estudios que verifiquen la estructura factorial del instrumento realizando una evaluación en 360° (Benton & Cashin's, 2012; Berk, 2009; 2014; Xu, 2012), donde además del criterio del alumnado se consideren otras valoraciones como son las autoevaluaciones del propio profesorado, informes de personas externas a la institución, compañeros de trabajo (Simona, Felicia, & Georgiana, 2014), etc. consiguiendo así una evaluación docente más integral. Este tipo de evaluación ofrece información valiosa desde diferentes perspectivas, lo que permite triangular la información y reorientar al docente con el fin de poner en marcha planes de mejora.

Además, se podrían considerar estudios longitudinales para analizar la validez de contenido y de constructo de esta escala ampliando la muestra de estudiantes y en los que se estimen modelos de regresión estructural que corroboren los resultados obtenidos.

REFERENCIAS

- Agencia Nacional de Evaluación de la Calidad (ANECA) (2007). *Programa Audit. Guía para el diseño de sistemas de garantía interna de calidad de la formación universitaria*. Retrieved from http://www.aneca.es/var/media/166338/audit_doc01_guidiseno_070621.pdf
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411-423. doi:10.1037/0033-2909.103.3.411
- Arranz, P. (2007). *Los sistemas de garantía de calidad en la educación superior en España. Propuesta de un modelo de acreditación para las titulaciones de grado en empresa* (Doctoral dissertation). Retrieved from <http://hdl.handle.net/10259/78>
- Balam, E., & Shannon, D. (2010). Student ratings of college teaching: A comparison of faculty and their students. *Assessment and Evaluation in Higher Education*, 35, 209-221. doi:10.1080/02602930902795901
- Barnes, D., Engelland, B., Matherne, C., Martin, W., Orgeron, C., Ring, J., et al. (2008). Developing a psychometrically sound measure of collegiate teaching proficiency. *College Student Journal*, 42, 199-213.
- Benton, S. L., & Cashin W. E. (2012). Student ratings of teaching: A summary of research and literature (IDEA Paper no. 50). Manhattan, KS: The IDEA Center. Retrieved from http://www.theideacenter.org/sites/default/files/idea-paper_50.pdf.
- Berk, R. A. (2009). Using the 360 degrees multisource feedback model to evaluate teaching and professionalism. *Medical Teacher*, 31(12), 1073-80. doi:10.3109/01421590802572775
- Berk, R. A. (2013). Top five flashpoints in the assessment of teaching effectiveness. *Medical Teacher*, 35, 15-26. doi:10.3109/0142159X.2012.732247
- Berk, R. A. (2014). Should Student Outcomes Be Used to Evaluate Teaching? *The Journal of Faculty Development*, 28(2), 87-96.
- Bol-Arreba, A., Sáiz-Manzanas, M. C., & Pérez-Mateos, M. (2013). Validación de una encuesta sobre la actividad docente en Educación Superior. *Aula Abierta*, 41(2), 45-54.
- Buendía Eisman, L. (coord.) (1996). Evaluación educativa (Monográfico). *Revista de Investigación Educativa*, 14(2).
- Byrne, B. M. (2001). *Structural equations modeling with Amos: Basic concepts, application, and programming*. Mahwah, NJ: Erlbaum.
- Carreras Barnés, J. (2005). El concepto de competencia y el diseño de planes de estudio a partir de perfiles competenciales. Dos cuestiones que hay que aclarar en el marco de las EEES. In *El Debate sobre las competencias a la enseñanza universitaria* (pp. 8-25). Barcelona: ICE UB.
- Casero, A. (2008). Propuesta de un cuestionario de evaluación de la calidad docente universitaria consensuado entre alumnos y profesores. *Revista de Investigación Educativa*, 26, 25-44.
- Cecchini, J. A., González, C., & Montero, J. (2007). Participación en el deporte y fairplay [Participation in sport and fair play]. *Psicothema*, 19, 57-64.
- Dodeen, H. (2013). Validity, Reliability, and Potential Bias of Short Forms of Students' Evaluation of Teaching: The Case of UAE University. *Educational Assessment*, 18(4), 235-250. doi:10.1080/10627197.2013.846670
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: a second course* (pp. 269-314). Greenwich: Information Age Publishing.
- García Ramos, J. M., & Congosto, E. (2000). Evaluación y Calidad del Profesorado. In T. González Ramírez (Ed.), *Evaluación y Gestión de la Calidad Educativa. Un Enfoque Metodológico* (pp. 127-157). Málaga: Aljibe.
- Gargallo, B. (2008). Estilos de docencia y evaluación de los profesores universitarios y su influencia sobre los modos de aprendizaje. *Revista Española de Pedagogía*, 66(241), 425-446.
- González, I., & López, I. (2010). Validación y propuesta de un modelo de indicadores de evaluación de la calidad en el la universidad. *Revista Iberoamericana de Educación*, 53(6), 1-13.
- González López, I., & López Cámara, A. (2010). Sentando las bases para la construcción de un modelo de evaluación a las competencias docentes del profesorado universitario. *Revista de Investigación Educativa*, 28(2), 403-423.
- Gursoy, D., & Umbreit, W. T. (2005). Exploring students' evaluations of teaching effectiveness: What factors are important? *Journal of Hospitality and Tourism Research*, 29(1), 91-109. doi:10.1177/1096348004268197
- Hattie, J. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Oxford, UK: Routledge.
- Keeley, J., Furr, R. M., & Buskist, W. (2010). Differentiating psychology students' perceptions of teachers using the Teacher Behavior Checklist. *Teaching of Psychology*, 37, 16-20. doi:10.1080/00986280903426282
- Kember, D., & Leung, D. (2008). Establishing the validity and reliability of course evaluation questionnaires. *Assessment & Evaluation in Higher Education*, 33, 341-353. doi:10.1080/02602930701563070
- Khong, T. L. (2014). The Validity and Reliability of the Student Evaluation of Teaching: A case in a Private Higher Educational Institution in Malaysia. *International Journal for Innovation Education and Research*, 2(9), 57-63.
- Kuzmanovic, M., Savic, G., Gusavac, B. A., Makajic-Nikolic, D., & Panic, B. (2013). A Conjoint-based approach to student evaluations of teaching performance. *Expert Systems with Applications*, 40(10), 4083-4089. doi:10.1016/j.esw.2013.01.039
- Leary, H., Walker, A., Shelton, B. E., & Harrison, F. M. (2013). Exploring the Relationships Between Tutor Background, Tutor Training, and Student Learning: A Problem-based Learning Meta-Analysis. *Interdisciplinary Journal of Problem-based Learning*, 7(1), 40-66. doi:10.7771/1541-5015.1331
- Leunga, X. Y., Jiangb, L., & Busserb, J. (2013). Online student evaluations of hospitality professors: A cross-cultural comparison. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 12(1), 36-46. doi:10.1016/j.jhlste.2012.10.001
- López-Barajas, D., & Carrascosa, J. (2005). La evaluación de la docencia universitaria. Dimensiones y variables más relevantes. *Revista de Investigación Educativa*, 23(1), 57-84.
- López Ruiz, J. I. (2011). Un giro copernicano en la enseñanza universitaria: formación por competencias. *Revista de Educación*, 356, 279-301. doi:10.4438/1988-592X-RE-2010-356-040
- Lukas, J. F., Santiago, K., Etxeberria, J., & Lizasoain, L. (2014). Adapting to the European Higher Education Area a questionnaire on student opinion about the teaching of lecturers. *Revista Electrónica de Investigación y Evaluación Educativa*, 20(1). doi:10.7203/relieve.20.1.3812
- Marqués, M. (2008, September). Una experiencia madura de innovación educativa. In *Jornadas Nacionales sobre Estudios Universitarios de los Proyectos de Convergencia a la Realidad de los Nuevos Títulos*. Universidad Jaume I.

- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-288. doi:10.1016/0883-0355(87)90001-2
- Marsh, H. W., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical self-description questionnaire: psychometric properties and a multi-trait-multimethod analysis of relations to existing instruments. *Journal of Sport and Exercise Psychology*, 16, 270-305.
- Moore, S., Walsh, G., & Rísquez, A. (2012). *Estrategias eficaces para enseñar en la Universidad*. Madrid: Narcea.
- Mortelmans, D., & Spoorén, P. (2009). A revalidation of the SET37-questionnaire for student evaluations of teaching. *Educational Studies*, 35, 547-552. doi:10.1080/03055690902880299
- Muñiz, J., & Bartram, D. (2007). Improving international tests and testing. *European Psychologist*, 12, 206-219. doi:10.1027/1016-9040.12.3.206
- Muñiz, J., & Fonseca-Pedrero, E. (2008). Construcción de instrumentos de medida para la evaluación universitaria. *Revista de Investigación en Educación*, 5, 13-25.
- Muñoz Cantero, J. M., Ríos de Deus, M. P., & Abalde, E. (2002). Evaluación Docente vs. Evaluación de la Calidad. (RELIEVE). *Revista Electrónica de Investigación y Evaluación Educativa*, 8(2), 103-134.
- Ntoumanis, N. (2001). A self-determination approach to the understanding of motivation in physical education. *British Journal of Educational Psychology*, 71, 225-242. doi:10.1348/000709901158497
- Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. T. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality & Quantity*, 43, 197-209. doi:10.1007/s11135-007-9112-4
- Padilla, M. T., & Gil, J. (2008). La evaluación orientada al aprendizaje en la Educación Superior: condiciones y estrategias para su aplicación en la docencia universitaria. *Revista Española de Pedagogía*, 66(241), 467-486.
- Pallisera, M., Fullana, J., Planas, A., & del Valle, A. (2010). La adaptación al espacio europeo de educación superior en España. Los cambios/retos que implica la enseñanza basada en competencias y orientaciones para responder a ellos. *Revista Iberoamericana de Educación*, 52(4), 1-13.
- Pozo, C., Giménez, M. L., & Bretones, B. (2009). La evaluación de la calidad docente en el nuevo marco del EEES. Un estudio sobre la encuesta de opinión del Programa DOCENTIA-ANDALUCÍA. *Educación XXI*, 11, 43-64.
- Pozo, C., Bretones, B., Martos, M. J., & Alonso, E. (2011). Evaluación de la actividad docente en el Espacio Europeo de Educación Superior: un estudio comparativo de indicadores de calidad en universidades europeas. *Revista Española de Pedagogía*, 248, 145-163.
- Salmerón, L. (2013). Actividades que promueven la transferencia de los aprendizajes: una revisión de la literatura. *Revista de Educación*, special issue 2013. doi:10.4438/1988-592X-RE-2013-EXT-253
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research*, 83(4), 598-642. doi:10.3102/0034654313496870
- Shukrie, R. (2011). Professor's performance for effective teaching (Kosovo case). *Procedia Social and Behavioral Sciences*, 12, 117-121. doi:10.1016/j.sbspro.2011.02.015
- Simona M., Felicia, S., & Georgiana, C. V. (2014). Content analysis regarding the multiple evaluation of the academic activity in CMU. *Global Journal of Guidance and Counselling*, 4(1), 25-31. Retrieved from <http://www.world-education-center.org/index.php/gjgc/article/view/3196>
- Tejedor, F. J., & García Valcárcel, A. (1996). Evaluación institucional en la Universidad. *Revista Galega de Psopedagogía*, 6, 101-146.
- Toland, M. D., & De Ayala, R. J. (2005). A multilevel factor analysis of students' evaluations of teaching. *Educational and Psychological Measurement*, 65, 272-296. doi:10.1177/0013164404268667
- Vansteenkiste, M., Soenens, B., Sierens, E., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology*, 101, 671-688. doi:10.1037/a0015083
- Xu, Y. (2012). Developing a comprehensive teaching evaluation system for foundation courses with enhanced validity and reliability. *Educational Technology Research and Development*, 60(5), 821-837. doi:10.1007/s11423-012-9240-y

v MIECEEEES: Modelo de Indicadores de evaluación de la calidad en el Espacio Europeo de Educación Superior. Financiado por el Ministerio de Educación y Ciencia, Dirección General de Universidades. Código: EA2006-0116

Con el fin de llegar a un mayor número de lectores, NAER ofrece traducciones al español de sus artículos originales en inglés. Sin embargo, **este artículo en español no es el artículo original sino únicamente su traducción**. Si quiere citar este artículo por favor consulte el artículo original en inglés y utilice la paginación del mismo en sus citas. Gracias.

i ANECA: Agencia Nacional de Evaluación de la Calidad y Acreditación

ii ACSUG: Axencia para a Calidade do Sistema Universitario de Galicia

iii AVAP: Agencia Valenciana d'Avaluació i Prospectiva

iv AQU: Agència per a la Qualitat del Sistema Universitari de Catalunya